

The Effect of Context on Distributed Practice

A Senior Honors Thesis

Presented in partial fulfillment of the requirements for graduation  
*with honors research distinction* in Psychology in the undergraduate colleges of  
The Ohio State University

by

Zachary Adolph Niese

The Ohio State University  
Spring 2013

Project Advisor: Dr. Per B. Sederberg, Department of Psychology

### **ABSTRACT**

One of the most robust effects in human memory literature is that of distributed practice: memory is better for a twice-presented item when there is an interval of time between the presentations. However, while the traditional explanation for this effect focuses on retrieval cues at recall, many current theories of memory place a heavy focus on encoding efficacy while learning. Thus, this project aims to explain the distributed practice effect in terms of encoding efficacy via an interaction of repetition priming (seeing an item twice in a row decreases encoding of the item and lowers memory for it) and an individual's ability to utilize their temporal context to make predictions. Participants were shown lists of stimuli in which "target" items were repeated over short and long lags. Additionally, before the second presentations of half of the targets, the same two preceding stimuli before the first presentation were shown (repeated context) while the other half were preceded by novel stimuli (novel context). Finally, participants were given a recognition task consisting of target stimuli and lures. We ran three studies, the first of which used medium-frequency words while the second and third used unfamiliar faces. Neither of the studies that used unfamiliar faces found significant results, presumably because the participants had no prior representation of the faces, which altered the encoding experience. However, we did find significant effects when using words. First, participants showed greater repetition priming when the context was repeated, resulting in poorer performance during the recognition task. Additionally, there appeared to be an interaction in which target words presented under novel context conditions began to suffer a detriment in performance over a longer lag while words presented under the repeated context condition showed no change. Thus, we suggest that although repeating the temporal context results in repetition priming and worse performance initially, it also makes individuals more likely to recognize that they saw the word before, allowing the spacing effect to persist over a longer lag.

## **THE EFFECT OF CONTEXT ON DISTRIBUTED PRACTICE**

On the whole, human memory is a highly sophisticated and impressive system that allows individuals to encode, store, and retrieve memories in a highly efficient manner. However, human memory is far from perfect, and we are simply incapable of remembering everything we experience and learn. Thus, an overarching aim in the psychological literature on memory is to discover why certain conditions give rise to successful memory storage and retrieval while others fail.

Although there are a multitude of factors influencing the memory system's efficacy, one of the most robust and well documented effects in the memory literature is that of distributed practice. The distributed practice effect contends that subsequent memory for a stimulus is enhanced when the stimulus is presented twice with an intervening lag of other stimuli or time between the presentations (known as spaced learning) as compared to a stimulus that is simply presented twice in a row (massed learning) (for a full review, see Cepeda et al, 2006). Numerous studies on list learning have found that participants are more likely to remember a stimulus that was presented under spaced learning conditions versus a stimulus presented under massed learning conditions, and although this effect has been replicated and studied a myriad of ways and times (Donovan & Radosevich, 1999), little work has actually been done to examine the driving force behind the effect itself.

The traditional explanation for the distributed practice effect, known as the encoding variability hypothesis, purports that spaced learning shows a boost in memory performance because the stimulus is linked to more retrieval cues (the words immediately preceding each presentation of the stimulus) than a stimulus that is presented under massed learning conditions (Melton, 1970). For example, in a list such as “chair, cat, fan, pencil, pencil, bottle, cat”, the word “cat” (spaced learning) is more likely to be remembered than “pencil” (massed learning)

## Context and Distributed Practice

because the word cat is associated with two words that will cue its retrieval (chair and bottle) whereas pencil is only associated with one retrieval cue (fan). However, although this explanation may work to explain some of the distributed practice effect, it focuses almost entirely on retrieval processes and is incapable of explaining variability on different experiences during encoding (e.g. features of the object, overall context, and ability to predict an upcoming event). This is a major pitfall of the explanation because, as evinced by a recent study, the stimuli a person will subsequently remember in list learning can be accurately predicted by his/her neural activity during encoding (Cheea, Westphala, Goha, Grahama, & Song, 2003). Thus, while the encoding variability hypothesis places emphasis upon retrieval cues in predicting subsequent memory, it appears that there is a strong case for instead placing emphasis on encoding efficacy.

Alternative explanations of the distributed practice effect that focus on the importance of encoding efficacy on subsequent memory come from more recent research on repetition priming. Repetition priming is the phenomenon in which people tend to pay less attention to a stimulus that they are familiar with, thereby exerting a lower amount of effort during encoding and resulting in poorer subsequent memory for the stimulus (Wagner, Maril, & Schacter, 2000). That is, when a stimulus is presented twice in a row under massed learning conditions, people pay less attention to the second presentation and receive little benefit from the repetition. However, when a stimulus is repeated after an interval of time has passed or interceding stimuli have been presented (i.e. spaced learning), people forget some of the features of the stimulus from the first presentation, thereby decreasing repetition priming and causing them to spend more time and effort encoding the second presentation (as compared to massed learning). This increased encoding effort, in turn, enhances the encoding efficacy of the stimulus and produces a

subsequent boost in memory for the item, as evinced in studies on both the behavioral and neural levels (Xue et al, 2010).

Further, explanations of distributed practice that utilize repetition suppression have additional explanatory power over the traditional explanation because they can be used to understand variances within spaced learning such as the lag effect. The lag effect is the phenomenon within spaced learning in which a longer lag (i.e. number of items or amount of time between the first and second presentation of the item) typically produces greater memory than a short lag (Thios & D'Agostino, 1976). That is, when an item is presented twice with an intervening amount of items or time (i.e. spaced learning), the more items that are presented or time that passes between the two presentations, the more likely a person is to remember the stimulus. However, it is important to note that the contextual variability encoding hypothesis is unable to explain the lag effect because the number of retrieval cues (i.e. the words preceding each presentation) remains constant across both short and long lags. In contrast, repetition priming does explain this effect because, as more time passes, people begin to forget more and more details of the stimulus. This reduces the repetition priming they show when the item is repeated, resulting in more effortful encoding and better memory (Henson, Rylands, Ross, Vuilleumeir, & Rugg, 2004).

However, one issue with an explanation of the spacing effect that solely relies on repetition priming is that it focuses entirely on the individual items and fails to account for the role of context in memory and encoding efficacy, and as memory researchers have long been aware, the contextual experience surrounding a learning episode can have a strong impact on subsequent memory (Godden & Baddeley, 1975). More recently, Howard and Kahana's (2002) Temporal Context Model of memory posits that context is fundamentally important in

understanding human memory and proposes that it works by creating time-dependent contextual states that are bound to one another sequentially in time. In this way, human memory is modeled as a process of continually binding everything in the current temporal context to the one before it. Working from this model of memory, we expect that a more accurate and sophisticated explanation of the spacing effect should be able to account for the effect of context on modulating and impacting repetition priming.

Indeed, recent studies on repetition priming suggest this is the case. For example, one study found that when a stimulus was repeated in a list that did not typically contain repetitions (i.e. it was surprising and out of the ordinary to see an item presented twice), the amount of repetition attenuation (the neural correlate of repetition priming in which the brain's response to a repeated stimulus is measurably smaller) was reduced, resulting in enhanced encoding and better subsequent memory (Summerfield, Trittschuh, Monti, Mesulam, and Egner, 2008). Further, another study found that when a repeated image was preceded by the same two images as its first presentation (which would serve to reinstate the temporal context it was originally bound to), there was a greater amount of repetition attenuation for the image as compared to when it was preceded by two different images (Turke-Browne, Simon, & Sederberg, 2012). Therefore, it is clear that context plays an important role in the modulation of repetition priming, and may have a significant effect on the underlying mechanisms of distributed practice.

Although reinstating a temporal context has been shown to increase repetition priming and decrease memory performance, there may be instances in which reinstating a similar context increases memory by allowing people to make predictions about future outcomes (Schacter & Addis, 2009). Additionally, it is possible that the act of consciously making a prediction directs attentional focus and effort to the encoding experience, thereby making the memory stronger

(Dudukovic, DuBrow, & Wagner, 2009). This has important implications because it has been suggested that the power of prediction is a “universal principle of the brain” (Bar, 2009) that is critically important for processes ranging from basic survival (Klein, Cosmides, Tooby, & Chance, 2002) to complex social interactions (Mitchell, 2009).

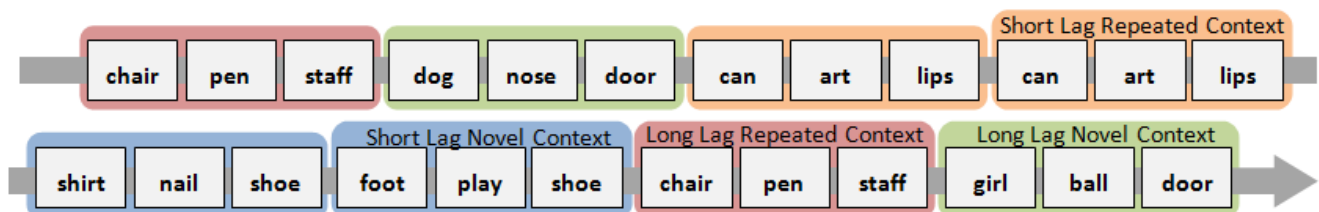
Therefore, the present project aims to look more closely at how reinstating the temporal context before the second presentation of a repeated stimulus modulates repetition priming while also interacting with the power of prediction to affect subsequent memory. More specifically, the studies involved subsequent memory recognition of target stimuli which were each presented twice across short and long lags (the number of stimuli between the two presentations) in a list. Additionally, for half of the target stimuli, the temporal context (defined as the two “context” stimuli presented before the target) was reinstated prior to the target’s repetition by also repeating the context stimuli before the target was presented, creating a “repeated context”. Thus, in the repeated context condition, participants were capable of using the reinstated temporal context to predict the upcoming stimulus. First, we hypothesize that this study design, consistent with prior research, will produce a general lag effect such that longer lags produce greater subsequent memory (Thios & D’Agostino, 1976). Additionally, we hypothesize that differences in context will interact with the lag effect. More specifically, we expect that when the context is repeated over a short lag, the target stimulus will be entirely expected and will require very little effort to predict, thereby increasing repetition priming and decreasing subsequent memory for the stimulus (as compared to targets in the “novel context” condition which were repeated without reinstating the original temporal context they were bound to). Conversely, when the lag is longer, it will be possible for individuals to recognize a repeated context, but predicting the target stimulus will require more effort and attention. Thus, this increase in attention and effort

required to make a prediction during encoding should create a boost in subsequent memory for the stimulus.

## STUDY ONE

### Introduction

Study One aimed to discover how reinstating the temporal context before a repeated word interacts with the lag effect in spaced learning. The study evaluated the performance of twice-presented target words (which only differed from other words in the list based on their placement) split into four groups: short-lag repeated context, short-lag novel context, long-lag repeated context, and long-lag novel context (see Figure 1). Context was defined as the two words preceding the target stimulus, and as such, in a repeated context, the second presentation of the target stimulus was preceded by the same two context words as the first presentation whereas in a novel context, the target stimuli was preceded by a different set of context words in the second presentation. Additionally, the amount of intervening stimuli between the two presentations of target stimuli was varied such that targets in the short-lag condition were presented with only two intervening stimuli (i.e. the context words) while target words in the long-lag condition were presented after lags of fifteen to twenty-one intervening words.



**Figure 1:** A basic depiction of the study design and illustration of the four conditions.



## **Methods**

### *Participants*

112 introductory psychology students at The Ohio State University participated in the study and were given course credit for their participation.

### *Materials*

Judgment task – participants were shown words one at a time from a list and were asked to decide if each word was living or nonliving as quickly and accurately as possible. The lists were constructed by randomly selecting words from a pool of average-frequency, semantically-unrelated words (Nelson, McEvoy, & Schreiber, 2004). Additionally, the lists were constructed such that they included an equal number of the four presentation types of the twice-presented target words (short-lag repeated context, short-lag novel context, long-lag repeated context, and long-lag novel context). Seven word lists were created per participant, each consisting of ninety-two words (sixty unique), sixteen of which were unique target words.

Recognition task – participants were shown words one at a time and were asked to decide if they had seen each word during the previous living/nonliving judgment task or not. Each recognition task consisted of the sixteen unique target words randomly shuffled with sixteen lures. There were seven recognition task lists, matching the seven word lists from the judgment task.

### *Design*

The study was constructed as a 2 (short-lag vs. long-lag) X 2 (repeated context vs. novel context) with-in subjects factorial design. Reaction times during the judgment task and performance on the recognition task was used as the dependent measures.

### *Procedure*

The study took slightly less than one hour to complete. After obtaining informed consent from the participants, the experimenter read the instructions for the judgment task (without informing the participants that their memory for the words would be tested later), simply asking participants to make their judgments as quickly and accurately as possible. After completing a judgment list, a screen appeared informing participants to press a button when they were ready to begin the next judgment list. Once participants completed all seven judgment lists, a prompt on the screen informed them that their recognition memory would be tested for the words they just saw. The recognition task lists were presented in sequential order with their matching judgment task (i.e. the first recognition task contained target words from the first judgment list, the second recognition task contained target words from the second judgment list, etc.) Once participants completed the final recognition task list, they were debriefed and thanked for their time and participation.

### **Results**

The data were analyzed using linear mixed effects models with statistical processing software R (R Development Core Team, 2009) and its packages *lme4* (Bates & Maechler, 2009) and *language R* (Baayen, 2009; cf. Baayen, 2008). The benefit of analyzing the data using a linear mixed effects model is that it is capable of isolating and accounting for differences in performance caused by variables other than the experimentally manipulated factors. Specifically, the model is able to account for individual differences between subjects such as genetic, developmental, social, and environmental factors (Baayan, Davidson, & Bates, 2008) and differences among individual stimuli that impact performance (Judd, Westfall, & Kenny, 2012).

Of the 112 total participants, 19 were excluded from analysis for failing to perform above chance on the test phase of the experiment. Thus, data for 99 subjects were analyzed using a linear mixed effects model. The following model was used to fit the data:

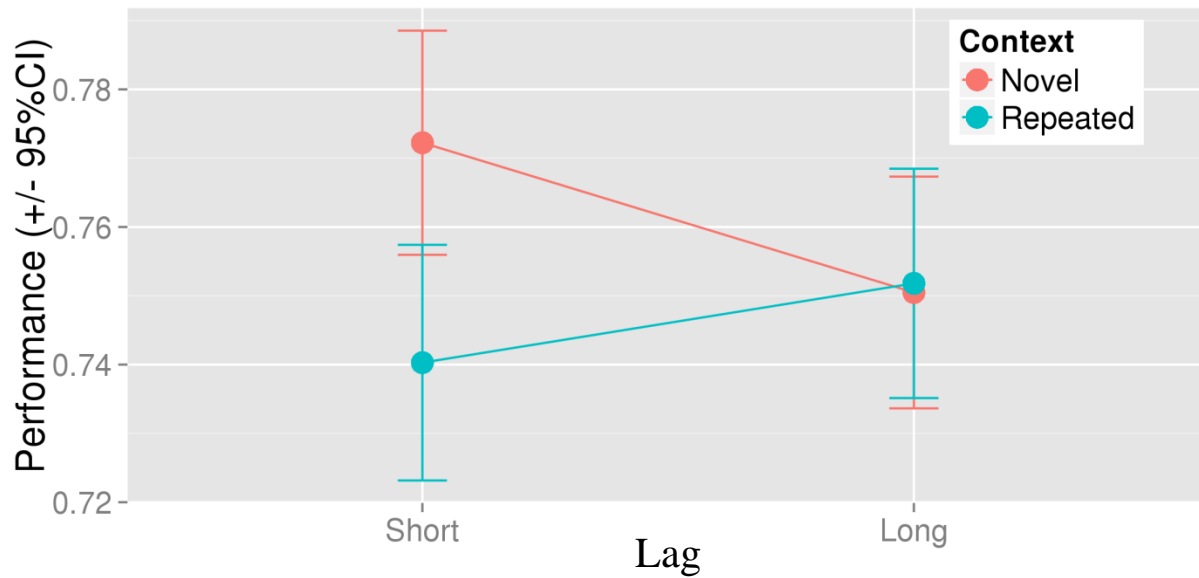
```
model.full = glmer(correct ~ lag * context + (1|subject) + (1|stimulus))
```

The above model attempts to fit the data by explaining performance at test (“correct”) as a function of the fixed effects for “lag” (i.e. short vs. long lag) and “context” (i.e. novel or repeated) while accounting for the random effects of individual subjects (1|subject) and stimuli (1|stimulus). To verify that the fixed effects in this model were adequately explaining differences in performance, the model was compared to the below null model that only included the random effects:

```
model.null = glmer(correct ~ (1|subject) + (1|stimulus))
```

We performed a likelihood ratio test to compare the fits of the full (model.full) and null (model.null) models using an ANOVA. We found that the difference was significant ( $p = .0497$ ), suggesting that the full model better fits and explains the data than the null model. Thus, looking at the fixed effects and their interaction, we found a significant main effect in which repeating the context yielded worse performance ( $p = .005$ ). Additionally, both a main effect in which longer lags yielded worse performance ( $p = .076$ ) and the interaction between context and lag ( $p = .053$ ) approached significance (see Figure 2).

A deeper analysis of this interaction looked at each half of the data split by context and found that there was no effect of lag condition for the repeated context ( $p = .375$ ), but there was an effect that approached significance in which novel contexts yielded worse performance over long lags ( $p = .074$ ).



**Figure 2:** A graph depicting the percentage of recognized target words across short and long lags within repeated and novel context conditions in Study One.

### *Reaction Time Analysis*

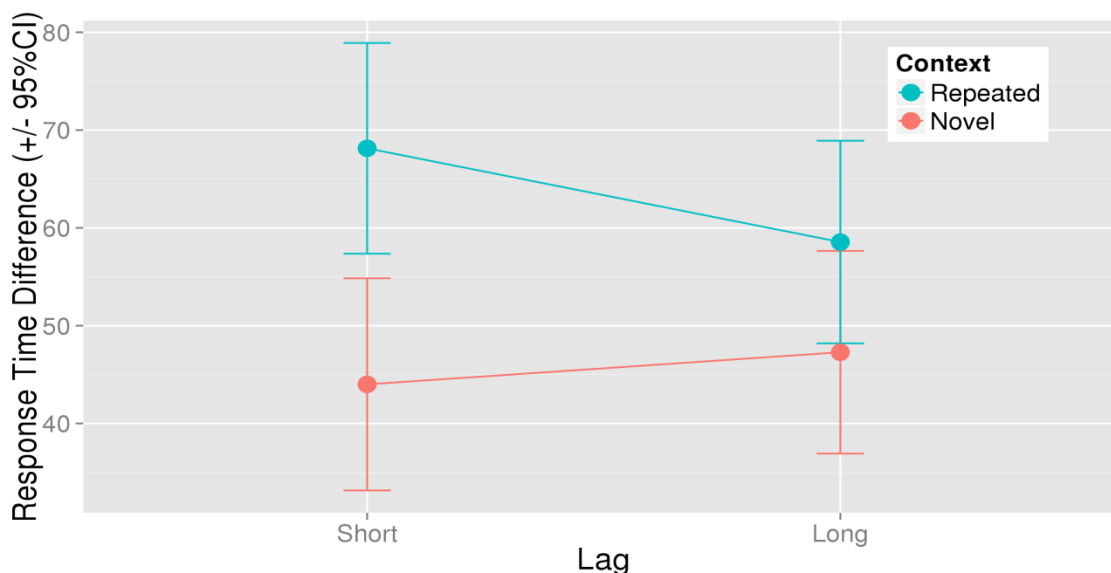
In addition to the performance analysis, we carried out an analysis on reaction times during the living/nonliving judgment task. Specifically, we were interested in analyzing to what extent participants were primed and/or expecting the second presentation of a stimuli by calculating how much faster they made the living/nonliving judgment for the second presentation (as compared to the first presentation). The following model attempts to capture these differences in reaction times as a function of the fixed effects of lag, context, and their interaction:

```
model.priming = lmer(judgment_rt_diff ~ lag * context + (1|subject) + (1|stimulus))
```

This model attempts to account for differences in “judgment\_rt\_diff” (the reaction time for the second presentation of an item subtracted from the reaction time for the first presentation). To verify that the fixed effects in this model were adequately explaining differences in reaction times, the model was compared to the below null model that only included the random effects:

```
model.priming.null = lmer(judgment_rt_diff ~ (1|subject) + (1|stimulus))
```

We performed a likelihood ratio test to compare the fits of the full (model.priming) and null (model.priming.null) models using an ANOVA. We found that the difference was significant ( $p = .0037$ ), suggesting that the full model better fits and explains the data than the null model. This analysis yields t-values, so we also ran a Markov chain Monte Carlo (MCMC) to convert the results into more interpretable p-values. Once again, we found a significant main effect for context ( $p = .001$ ). However, there was no significant main effect for lag ( $p = .697$ ) and no significant interaction ( $p = .237$ ). As can be seen in Figure 3, repeating the context resulted in a greater difference in reaction time (i.e. participants made the judgment faster the second time an item was presented when its context was repeated).



**Figure 3:** A graph depicting differences in reaction times for target words during the judgment task across short and long lags within repeated and novel context conditions in Study One.

## Discussion

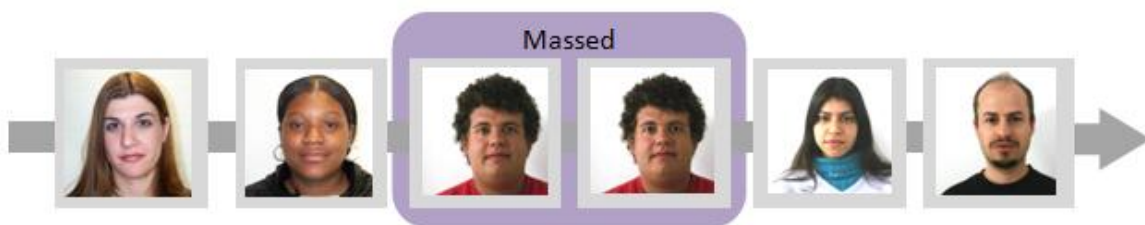
The performance results, as depicted in Figure 2, show the main effect in which repeating the context before the second presentation yielded lower subsequent memory for the target word. When combined with the reaction time analysis depicted in Figure 3, this suggests that the

repeated context primed participants to expect the target stimulus across both long and short lags, resulting in greater repetition priming and lower subsequent memory. Although the main effect for context seems consistent and fairly easy to interpret, the interaction between context and condition is slightly less clear. As seen in Figure 1, there appears to be a difference between repeated and novel contexts for a short lag, but this difference becomes non-existent over the long lag. Additionally, it is interesting to note that the reason for this interaction is not that the repeated context showed a significant boost over the long lag (as originally hypothesized) but seems to be driven by the novel context condition suffering a detriment over the long lag, contrary to what previous studies on lag effect would predict (Thios & D'Agostino, 1976).

## STUDY TWO

### Introduction

Study Two was designed with the aim of replicating, extending, and clarifying the results from the first study. First, because Study One seemed to find results that contradicted the expected lag effect, we added a massed condition (in which the second presentation of a stimuli immediately followed the first presentation) to ensure that, at the very least, the study was able to replicate the more established and powerful distributed practice effect (see Figure 4). Additionally, we were interested to see if the findings from Study One could be replicated and extended to include stimuli other than words, so Study Two was composed of images of faces from an online database of adults aged 18-50 (Minear and Park, 2004).



**Figure 4:** An illustration of the massed condition and examples of non-familiar faces as stimuli.

## **Methods**

### *Participants*

72 introductory psychology students at The Ohio State University participated in the study and were given course credit for their participation.

### *Materials*

Judgment Task – participants were once again shown stimuli from a list once at a time. However, unlike study one, the stimuli were drawn from a pool of ordinary faces and participants were asked to decide if each face was either male or female as quickly and accurately as possible. Each list included an equal number of male and female faces. The list was once again constructed to contain an equal number of the four conditions of target images (short-lag repeated context, short-lag novel context, long-lag repeated context, and long-lag novel context). Additionally, a fifth “massed” condition was included in which the target image was simply repeated twice in a row with no interceding stimuli. There were five judgment-task lists per participant, each consisting of 104 images of faces (sixty-four unique), twenty of which were unique target images.

Recognition task - participants were shown images of faces one at a time and were asked to decide if they had seen each word during the previous male/female judgment task or not. Each recognition task consisted of the twenty unique target words randomly shuffled with twenty lures. There were five recognition task lists, matching the five lists from the judgment task.

### *Design*

The study was constructed as a 2 (short-lag vs. long-lag) X 2 (repeated context vs. novel context) with-in subjects factorial design. Additionally, the massed condition was compared to the other lag conditions to verify that the general spacing effect was being replicated. Reaction

times during the judgment task and performance on the recognition task was used as the dependent measures.

### *Procedure*

The procedure was the same as Study One, consisting of completing the five judgment task lists followed by completing the five recognition task lists.

### **Results**

Of the 72 total participants, 26 were excluded from analysis for failing to perform above chance on the test phase of the experiment. Thus, data for 46 subjects were analyzed using a linear mixed effects model. First, we analyzed if the massed condition showed the expected detriment in performance as compared to the short and long lag spaced conditions using the following model to fit the data:

$$\text{model.massed} = \text{glmer}(\text{correct} \sim \text{spacing} + (1|\text{subject}) + (1|\text{stimulus}))$$

To check if the fixed effect of spacing was adequately explaining differences in performance, the model was compared to the below null model:

$$\text{model.massed.null} = \text{glmer}(\text{correct} \sim (1|\text{subject}) + (1|\text{stimulus}))$$

We performed a likelihood ratio test to compare fits of model.massed and model.massed.null using an ANOVA. We found that the difference was significant ( $p < .00001$ ). Thus, we interpreted the p-values in model.massed and found that the massed condition showed significantly worse performance than both the short and long lag spaced conditions ( $p < .0001$  for both).

After verifying that the study replicated the distributed practice effect, targets in the massed condition were excluded from subsequent analyses. The following model was used to fit the data:

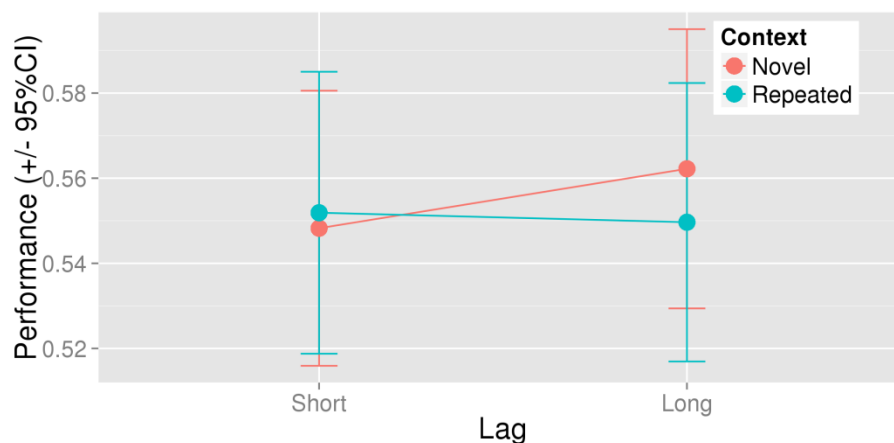


```
model.full = glmer(correct ~ lag * context + (1|subject) + (1|stimulus))
```

To check if the fixed effects in this model were adequately explaining differences in performance, the model was compared to the null model below that only included the random effects:

```
model.null = glmer(correct ~ (1|subject) + (1|stimulus))
```

We performed a likelihood ratio test to compare the fits of the full (model.full) and null (model.null) models using an ANOVA. We found that the difference was not significant ( $p = .985$ ), suggesting that the fixed effects in the full model (i.e. the independent variables) were not accounting for differences in performance as can be seen in Figure 5.



**Figure 5:** A graph depicting the percentage of recognized target words across short and long lags within repeated and novel context conditions in Study Two.

### *Reaction Time Analysis*

Although the performance analyses yielded no significant effects for the independent variables or their interaction, we carried out an analysis on reaction times during the male/female judgment task to see if there were priming differences during encoding. Once again, the

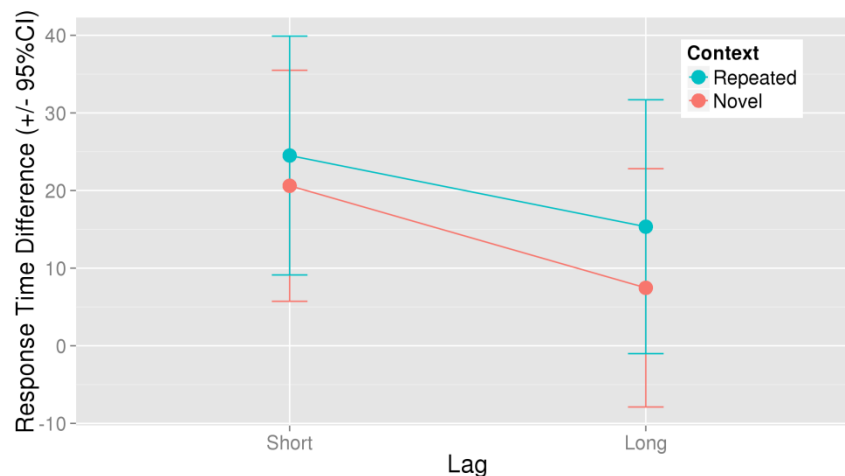
following model attempts to capture the difference in reaction times between the first and second presentations of a stimulus as a function of the fixed effects of lag, context, and their interaction:

```
model.priming = lmer(judgment_rt_diff ~ lag * context + (1|subject) + (1|stimulus))
```

To check if the fixed effects in this model were adequately explaining differences in reaction times, the model was compared to the below null model that only included the random effects:

```
model.priming.null = lmer(judgment_rt_diff ~ (1|subject) + (1|stimulus))
```

We performed a likelihood ratio test to compare the fits of the full (model.priming) and null (model.priming.null) models using an ANOVA. We found that the difference was not significant ( $p = .423$ ), suggesting that the fixed effects in the priming model (i.e. the independent variables) were not accounting for differences in reaction times.



**Figure 6:** A graph depicting differences in reaction times for target words during the judgment task across short and long lags within repeated and novel context conditions in Study Two.

## Discussion

As can be seen from the results depicted in Figures 5 and 6, Study Two failed to find any significant differences in performance or reaction times for context, lag condition, or their interaction (beyond the general detriment for massed learning). However, there are a number of

potential reasons why this study failed to replicate Study One's results. First, recognition memory for faces is apparently much more difficult than for words (nearly a third of the total participants were excluded for performing at chance and of those kept, performance was still barely above 50%). Additionally, reaction times may not have shown a difference based on context or condition simply because it is much harder to rehearse images than words, which could have resulted in attenuated the repetition priming effect for repeated context.

### **STUDY THREE**

#### **Introduction**

Study Three was designed with the aim of addressing the problems with Study Two. The primary difference in Study Three's design is that participants completed the recognition task for each list immediately following each judgment task. We expect this change in design to affect the results in a few ways. First, by testing participants' memory immediately following study, it is likely that performance will increase overall for the individuals that are actually putting effort into the task. Second, this design switches the learning from incidental (i.e. participants were not told to remember the stimuli but recognized them at test anyway) to intentional, which may motivate participants to contribute more effort during encoding, increase the likelihood that they utilize the predictable context, and more distinctly pronounce the effect of context. Finally, previous work on lag effects have shown that they are much more fickle than the basic spacing effect and are heavily dependent upon the relationship between lag length and retention interval (Cepeda et.al., 2006). Thus, because the differences in lag were relatively small (fifteen to twenty-one items), shortening the retention interval was expected to pronounce the potential effect of lag condition and its interaction with context.

## **Methods**

### *Participants*

101 introductory psychology students at The Ohio State University participated in the study and were given course credit for their participation.

### *Materials*

Judgment task – participants were once again shown images of faces one at a time from a list and were asked to decide if each face was male or female. The list was once again constructed to contain an equal number of the five conditions of target images (short-lag repeated context, short-lag novel context, long-lag repeated context, long-lag novel context, and massed). There were five judgment-task lists per participant, each consisting of 104 images of faces (sixty-four unique), twenty of which were unique target images.

Recognition task - participants were shown images of faces one at a time and were asked to decide if they had seen each word during the previous male/female judgment task or not. Each recognition task consisted of the twenty unique target words and twenty lures. Unlike the recognition task lists in the previous two studies, the task was constructed so that items from the first half of the targets from the matching judgment task were tested in the first half of the recognition task (in random order with lures) and the second half were tested second. There were five recognition task lists, matching the five lists from the judgment task.

### *Design*

The study was constructed as a 2 (short-lag vs. long-lag) X 2 (repeated context vs. novel context) with-in subjects factorial design. Additionally, the massed condition was compared to the other lag conditions to verify that the general spacing effect was being replicated. Reaction

times during the judgment task and performance on the recognition task was used as the dependent measures.

### *Procedure*

The study took slightly less than one hour to complete. After obtaining informed consent from the participants, the experimenter read the instructions for the judgment task, asking participants to make their judgments as quickly and accurately as possible. However, unlike the previous two studies, participants were also informed that they should try to commit each face to memory because there would be a recognition task to follow. After completing the first judgment task, a screen gave participants instructions for the recognition task asking them to indicate whether they had seen each face in the judgment task they had just completed or not. After the recognition task, participants began another judgment list, followed by its recognition task. This process repeated until all five judgment and recognition lists were completed, at which point participants were debriefed and thanked for their time.

### **Results**

Of the 101 total participants, 10 were excluded from analysis for failing to perform above chance on the test phase of the experiment. Thus, data for 91 subjects were analyzed using a linear mixed effects model. Again, we began by analyzing if the massed condition showed the expected detriment in performance as compared to the short and long lag spaced conditions using the following model to fit the data:

$$\text{model.massed} = \text{glmer}(\text{correct} \sim \text{spacing} + (1|\text{subject}) + (1|\text{stimulus}))$$

To check if the fixed effect of spacing was adequately explaining differences in performance, the model was compared to the below null model:

$$\text{model.massed.null} = \text{glmer}(\text{correct} \sim (1|\text{subject}) + (1|\text{stimulus}))$$

## Context and Distributed Practice

We performed a likelihood ratio test to compare fits of model.massed and model.massed.null using an ANOVA. We found that the difference was significant ( $p < .00001$ ). Thus, we interpreted the p-values in model.massed and found that the massed condition showed significantly worse performance than both the short and long lag spaced conditions ( $p < .0001$  for both).

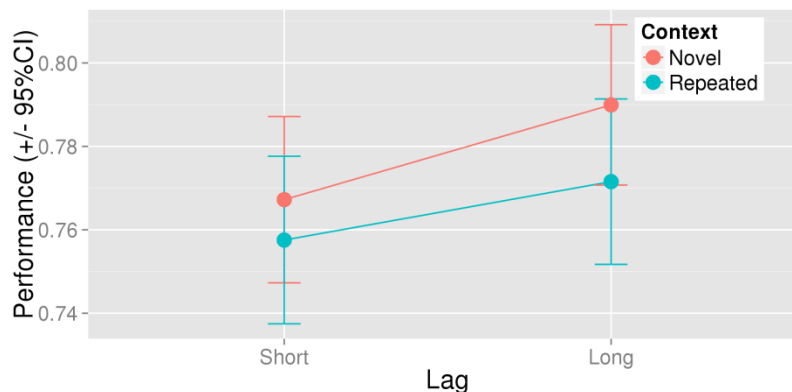
After verifying that the study replicated the distributed practice effect, targets in the massed condition were excluded from subsequent analyses. The following model was used to fit the data:

```
model.full = glmer(correct ~ lag * context + (1|subject) + (1|stimulus))
```

To check if the fixed effects in this model were adequately explaining differences in performance, the model was compared to the null model below that only included the random effects:

```
model.null = glmer(correct ~ (1|subject) + (1|stimulus))
```

We performed a likelihood ratio test to compare the fits of the full (model.full) and null (model.null) models using an ANOVA. We found that the difference was not significant ( $p = .142$ ), suggesting that the fixed effects in the full model (i.e. the independent variables) were not accounting for differences in performance.



**Figure 7:** A graph depicting the percentage of recognized target words across short and long lags within repeated and novel context conditions in Study Three.

*Reaction Time Analysis*

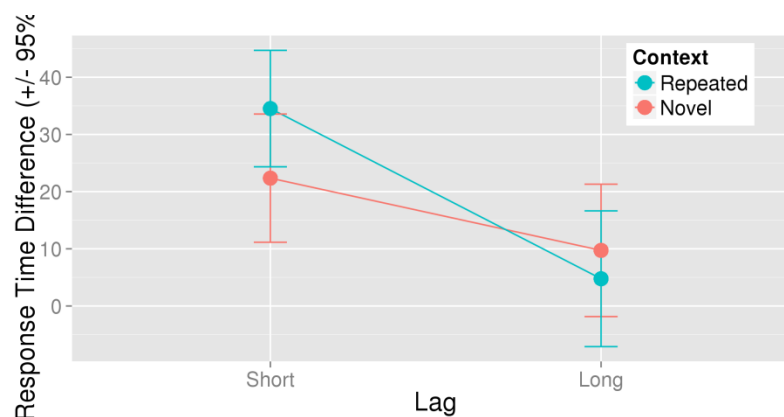
Although the performance analyses yielded no significant effects for the independent variables or their interaction, we carried out an analysis on reaction times during the male/female judgment task to see if there were priming differences during encoding. Once again, the following model attempts to capture the difference in reaction times between the first and second presentations of a stimulus as a function of the fixed effects of condition, context, and their interaction:

```
model.priming = lmer(judgment_rt_diff ~ lag * context + (1|subject) + (1|stimulus))
```

To check if the fixed effects in this model were adequately explaining differences in reaction times, the model was compared to the below null model that only included the random effects:

```
model.priming.null = lmer(judgment_rt_diff ~ (1|subject) + (1|stimulus))
```

We performed a likelihood ratio test to compare the fits of the full (model.priming) and null (model.priming.null) models using an ANOVA. We found that the difference was significant ( $p < .001$ ), suggesting that the fixed effects in the priming model (i.e. the independent variables) were accounting for differences in reaction times. However, when we used the MCMC to convert the t-values yielded by the full model into p-values, there were no significant results for lag condition ( $p = .103$ ), context ( $p = .134$ ), or their interaction ( $p = .136$ ).



**Figure 8:** A graph depicting differences in reaction times for target words during the judgment task across short and long lags within repeated and novel context conditions in Study Three.

## **Discussion**

Although Study Three succeeded in improving participants' performance overall, it also failed to find any significant differences in performance or reaction times for context, lag condition, or their interaction (beyond the detriment for massed learning). However, when looking at the performance plot in Figure 7, we can see that there seems to be a trend in performance in which people did slightly worse in repeated contexts and slightly better across long lags. Additionally, by looking at the reaction time plots in Figure 8, there seems to be corollary trends in which people showed a slight repetition priming effect for repeated contexts (in the short lag condition) and greater repetition priming for short lags overall. However, as noted above, these are not significant differences and the interpretation should be seen as suggestive at best.

## **GENERAL DISCUSSION**

As discussed above, consistent with the findings by Xue et. al. (2011), Study Two and Three found a significant boost in performance for facial recognition of spaced items (i.e. short and long lags) as compared to massed items, but both failed to produce a lag effect and showed no effect of context or their interaction. However, the lack of effect may have been due to the use of unfamiliar faces as stimuli. One reason that unfamiliar faces may have failed to produce a lag effect is because, unlike a study using words, participants entered the experiment with no prior representation of the stimuli. Thus, the experience at encoding would be very different (particularly for the first presentation) because rather than activating a previously-stored representation, the participant had to expend energy constructing a completely novel representation. Indeed, previous studies have shown that while words produce lag effects, novel



stimuli with no prior representation before the study (such as unfamiliar faces and non-words) typically do not produce lag effects (Bentin & Moscovitch, 1988). Another issue with using unfamiliar faces as stimuli is that although there is evidence that people use subverbal rehearsal to remember digits, letters, and words (Baddeley, 1986; Baddeley & Gathercole, 1998), there is no evidence that this process occurs (or is even possible) when trying to remember unfamiliar faces. Thus, even if participants were able to recognize when the context was repeated, it could be very difficult to consciously predict the upcoming face without an associated label or name to subverbally rehearse. Therefore, while the results from Study Two and Study Three did not support the original hypothesis, there is reason to believe that this was a result of using unfamiliar faces as stimuli.

Although Study Two and Study Three failed to yield significant findings, the results from Study One (which used medium-frequency words as stimuli) warrant further discussion. First, the findings show that participants performed worse when the context before the first and second presentations of target items was repeated. In addition, the reaction time analysis showed that participants made the living/nonliving judgment significantly faster when the context was repeated. This suggests that the differences in performance are attributable to repetition priming, which is in line with previous fMRI studies that have suggested repetition priming serves to attenuate neural activation during encoding and thereby result in worse subsequent memory (Callan & Schweighofer, 2009; Wagner, Koutstaal, Maril, Schacter & Buckner, 2000).

In addition to the main effect of context, the results of Study One suggest a potential interaction between context and lag condition. Specifically, although there appears to be a difference between novel and repeated contexts within a short lag, the difference disappears over the long lag. However, this interaction is difficult to interpret, particularly because there was not

a boost in performance for long lags within the novel condition as was expected based on previous research (Thios & D'Agostino, 1976). Thus, it would be extremely beneficial to further investigate this interaction by running another study using words that utilized the design of Study Three in which the difference in lag length better matched retention interval- a critical component in maximizing the lag effect's efficacy (Cepeda et. al., 2006).

However, the current results suggests that the power of prediction did not interact with the lag effect as initially expected. That is, although it appears that participants were able to predict the upcoming target word (as evinced by their faster reaction times), this did not translate to the increased effort and attention in encoding that was expected to boost performance over long lags. One potential explanation for this is that the brain is constantly making subconscious predictions (Bar, 2007) and that learning is not facilitated by the act of making predictions but rather, by prediction error (Hollerman & Schultz, 1998). This explanation of prediction suggests that people are making subconscious predictions at every moment of their lives about what will come next (e.g. I will see my coworker at his desk once I step into the office) and it only catches our attention and enhances learning when there is an error in our predictions (that is, we notice when the coworker is sick and does not show up to work that day). Thus, although it seems plausible that participants were making predictions about the target words based on the temporal context, these predictions did not enhance subsequent memory because there was no prediction error.

Nonetheless, there are potential explanations for the observed interaction. First, in understanding the lack of a lag effect in the novel context, it is important to note that the lag effect is not unbounded and is posited to function as an inverted-u shape (Küpper-Tetzel & Erdfelder, 2011). That is, although increasing lag lengths boost performance at first, lags that

become too long result in the person forgetting the first presentation of the stimulus and result in poorer performance. Thus, it is possible that participants did not realize that some of the words in the novel-context long-lag condition were repeated because they forgot the first presentation, preventing them from receiving the benefit of spaced learning. This explanation could be tested by running a similar study that included a wider range of lag conditions. If this explanation proved veritable, the study should not only replicate the results from Study One but should also find that performance for target words with a novel context continues to decrease as the lag length increases.

Additionally, because this explanation relies on an inverted u-shaped function of the lag effect that is based on forgetting, it can be used to understand the findings in the repeated context conditions. Specifically, research on context-dependent memory (a boost in memory performance when a person's memory is tested in the same context as study) would suggest that by repeating the temporal context before a target increased the possibility, it was more likely that participants would remember the first presentation of the target stimulus (Godden & Baddeley, 1975). In this way, repeating the context before the target word increased the likelihood that it would be recognized as a repeated word and continue to benefit from the spacing effect. These results would be consistent with a study on paired association learning that found that the second word of pairs were less likely to be remembered across a short lag if the second presentation showed the exact same pairing (versus pairing the second word with a novel word), but the inverse was true across a long lag (i.e. repeating the same pair twice over a long lag boosted memory for the second word) (Madigan, 1969). Thus, by reinstating the target word's original temporal context (i.e. the two preceding words) in the repeated context condition in Study One, it became more likely that the participants would remember the first presentation of the target and

continue to benefit from the spacing effect (Kahana & Howard, 2005). Therefore, if we were to carry out a study including a variety of longer lags, we would expect that reinstating the temporal context prior to the second presentation of a stimulus would initially result in poorer performance over short lags (due to repetition priming) but would also serve as a buffer from forgetting as lags increased, thereby causing repeated-context target words to outperform target words in the novel context condition as their performance continued to decline.

In conclusion, these results are in accordance with previous research that suggest repetition priming plays a crucial role as an underlying mechanism of the spacing effect. However, we additionally propose that the temporal context is an important factor in modulating the effect. Specifically, these results suggest that although reinstating the temporal context increases repetition priming (and decreases performance), it also serves as a buffer to prevent the forgetting that causes the inverted-u shape function of the spacing effect (Cepeda, et. al. 2008). Thus, consistent with recent research on the spacing effect and in the memory literature in general, it is clear that the experience during encoding (which is heavily impacted by temporal context) is critical to memory formation.

## WORKS CITED

- Baayen, R.H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R.H., Davidson, D.J., Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.
- Baayen, R. H. (2009). *languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics"*. R package version 0.955.
- Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as language learning device. *Psychological review*, 105(1), 158.
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11, 280-289.
- Bar, M. (2009). Predictions: A universal principle in the operation of the human brain. *Phil. Trans. R. Soc. B*, 364, 1181-1182.
- Bates, D.M. & Maechler, M. (2009). *lme4: Linear mixed-effects models using Eigen and R*. R package version 0.999375-32.
- Bentin, S., & Moscovitch, M. (1988). The time course of repetition effects for words and unfamiliar faces. *Journal of Experimental Psychology General*, 117, 148-148.
- Callan, D. E., & Schweighofer, N. (2010). Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory. *Human brain mapping*, 31(4), 645-659.
- Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., & Rohrer, D. (2006). Distributed practice in Verbal recall tasks: a review and quantitative synthesis. *Psychological Bulletin*, 132, 354-380.
- Cepeda, N.J., Vul, E., Rohrer, D., Wixted, J.T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridge of optimal retention. *Psychological Science*, 19, 11, 1095-1102.
- Cheea, M.W., Westphala, C., Goha, J., Grahama, S., & Song, A.W. (2003). Word frequency and subsequent memory effects studied using event-related fMRI. *Neuroimage*, 20, 1042-1051.
- Donovan, J.J. & Radosevich, D.J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, 84, 5, 795-805.
- Dudukovic, N.M., DuBrow, S., & Wagner, A.D. (2009). Attention during memory retrieval enhances future remembering. *Memory and Cognition*, 37, 7, 953-961.

- Godden, D.R., & Baddeley, A.D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66, 325-332.
- Henson, R.N., Rylands, A., Ross, E., Vuilleumeir, P., & Rugg, M.D. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *Neuroimage*, 21,4, 1674-1689.
- Hollerman, J.R. & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, 1, 304-309.
- Howard, M.W. & Kahana, M.J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, 46, 3, 269-299.
- Judd, C.M., Westfall, J., & Kenny, D.A. (2012). Treating stimuli as a random factor in Social Psychology: A new comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103, 1, 54-69.
- Kahana, M.J. & Howard, M.W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*, 12, 1, 159-164.
- Klein, S., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: multiple systems, multiple functions. *Psychological Review*, 109, 306-329.
- Küpper-Tetzel, C.E. & Erdfelder, E. (2012). Encoding, maintenance, and retrieval processes in the lag effect: A multinomial processing tree analysis. *Memory*, 20(1), 37-47.
- Madigan, S.A. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, 8, 828-835.
- Melton, A.W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 9, 596-606.
- Minear, M. & Park, D.C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*. 36, 630-633.
- Mitchell, J.P. (2009). Inferences about mental states. *Phil. Trans. R. Soc. B*, 364, 1309-1316.
- Nelson, D.L., McEvoy, C.L., & Schreiber, T.A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods*, 36, 3, 402-407.
- R Development Core Team (2009) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Schacter, D.L. & Addis, D.R. (2009). On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philosophical Transactions of the Royal Society, Biological Sciences*, 364, 1521, 1245-1253.
- Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, 11,

1004-1006.

- Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, 15, 529–536.
- Turk-Browne, N. B., Simon, M. G., Sederberg, P. B. (2012). Scene representations in parahippocampal cortex depend on temporal context. *Journal of Neuroscience*, 32.
- Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., & Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex*, 10(12), 1176-11
- Wagner, A.D, Maril, A., & Schacter, D.L. (2000). Interactions between forms of memory: When priming hinders new episodic learning. *Journal of Cognitive Neuroscience*, 12, 52-60.
- Xue, G., Mei, L., Chen, C., Lu, Z., Poldrack, R., & Dong, Q. (2010). Facilitating memory for novel characters by reducing neural repetition suppression in the left fusiform cortex. *PLS ONE*, 5, e13204.